

Black hole formation from massive scalar field collapse in the Einstein-de Sitter universe

Sérgio M. C. V. Gonçalves

Theoretical Astrophysics, California Institute of Technology, Pasadena, California 91125

(February 7, 2008)

We study the spherically symmetric collapse of a real, minimally coupled, massive scalar field in an asymptotically Einstein-de Sitter spacetime background. By means of an eikonal approximation for the field and metric functions, we obtain a simple analytical criterion—involving the physical size and mass scales (the field’s inverse Compton wavelength and the spacetime gravitational mass) of the initial matter configuration—for generic (non-time-symmetric) initial data to collapse to a black hole. This analytical condition can then be used to place constraints on the initial primordial black hole spectrum, by considering spherical density perturbations that re-entered the horizon during an early matter-dominated phase that immediately followed inflation.

PACS numbers: 04.25.-g, 04.40.-b, 04.70.Bw, 98.80.-k.

I. INTRODUCTION

Until now no fundamental elementary spin-0 particle has been detected in accelerator experiments, even though its existence is predicted by the so far highly successful $SU(3) \otimes SU(2) \otimes U(1)$ Standard Model [1]. Despite this lack of knowledge, most models of early universe astroparticle physics include at least one such field, and their remarkable agreement with observational data (see e.g. [2] for a recent review) makes pertinent the question of whether primordial density fluctuations in the scalar field energy density distribution could have collapsed to form black holes.

In particular, several models of inflation predict the existence of a period dominated by the energy of a real massive scalar field, just after the end of the inflationary epoch [3–5]. During this period, the field can behave like dust and density inhomogeneities can undergo non-linear growth, which may lead to the formation of primordial black holes [6]. It is during this early matter dominated phase that black hole formation—via gravitational collapse—is the most abundant, and we should therefore expect important constraints on inflationary models from the overproduction of primordial black holes during this epoch.

In this paper, we address one aspect of this problem by obtaining analytical conditions for black hole formation during the early matter dominated phase; the constraining of the primordial black hole energy density spectrum and its implications for inflationary scenarios will be discussed in detail elsewhere [7].

Our model consists of a real, spherically symmetric massive scalar field, minimally coupled to gravity, evolving in an asymptotically Einstein-de Sitter (EdeS) spacetime background. The Einstein equations for gravity coupled to the massive scalar field are solved analytically, using an asymptotic expansion for the field (and metric functions), in terms of the field’s Compton wavelength,

μ^{-1} , to explore the large mass limit. In such a limit, the scalar field behaves like general inhomogeneous dust [8], described by the Tolman-Bondi metrics [9–12]. From generic initial data, spherical dust collapse always proceeds to a black hole* and thus, by imposing the condition that the eikonal approximation holds, we obtain a sufficient criterion for black hole formation from massive scalar field collapse.

This paper is organized as follows: Section II derives the field equations for the massive Einstein-Klein-Gordon system, which are then analytically solved using an asymptotic series (in μ^{-1}) for the field and metric functions. The large mass limit of the eikonal approximation is the Tolman-Bondi family of metrics, described in Sec. III. In Sec. IV, the gravitational collapse of spherical density perturbations in the EdeS universe—a particular case of the general Tolman-Bondi spacetimes—is analyzed in terms of initial data. In Sec. V, analytical conditions for the initial data to collapse to a black hole are obtained by enforcing the validity of the WKB approximation to second order in μ^{-1} . Section VI discusses how this condition for black hole formation might affect the constraining of the initial (at the early matter dominated phase) primordial black hole spectrum, and thus the inflationary models responsible for generating the initial density fluctuation spectrum. Section VII concludes with a summary and a brief discussion on avenues for future work.

Geometrized units, in which $G = c = 1$, are used throughout.

*Central naked singularities can form from regular initial data, but they are *not* generic: only one non-spacelike geodesic can escape from the singularity, which, although strong, is also massless [13].

II. THE MASSIVE EINSTEIN-KLEIN-GORDON SYSTEM

We consider a general spherically symmetric metric, written here in normal Gaussian coordinates $\{t, r, \theta, \varphi\}$:

$$ds^2 = -dt^2 + e^{-2\Lambda} dr^2 + R^2 d\Omega^2, \quad (1)$$

where $\Lambda = \Lambda(t, r)$, $R = R(t, r)$, and $d\Omega^2 = d\theta^2 + \sin^2 \theta d\varphi^2$ is the canonical metric of the unit 2-sphere.

The independent non-vanishing Einstein tensor components are

$$G_{tt} = R^{-2}[-Re^{2\Lambda}(2R'\Lambda' + 2R'' + R^{-1}R'^2) - 2\dot{R}\dot{\Lambda}R + 1 + \dot{R}^2], \quad (2)$$

$$G_{rt} = -2R^{-1}(\dot{R}' + R'\dot{\Lambda}), \quad (3)$$

$$G_{rr} = -R^{-2}[e^{-2\Lambda}(2\ddot{R}R + \dot{R}^2 + 1) - R'^2], \quad (4)$$

$$G_{\theta\theta} = \sin^{-2} \theta G_{\varphi\varphi} = R[\dot{R}\dot{\Lambda} + R'\Lambda'e^{2\Lambda} + R''e^{2\Lambda} - \ddot{R} + \ddot{\Lambda}R - \dot{\Lambda}^2 R], \quad (5)$$

where the overdot and prime denote partial differentiation with respect to t and r , respectively.

For the matter content, we consider a real minimally coupled scalar field ϕ of mass μ , governed by the Klein-Gordon equation:

$$(\square - \mu^2)\phi = 0. \quad (6)$$

With the spherically symmetric metric (1) we have

$$\ddot{\phi} - e^{2\Lambda}\phi'' + \mu^2\phi + (\dot{\Lambda} + 2R^{-1}\dot{R})\dot{\phi} - e^{2\Lambda}(\Lambda' - 2R^{-1}R')\phi' = 0. \quad (7)$$

The stress-energy tensor of the scalar field is given by

$$T_{ab} = \nabla_a \phi \nabla_b \phi - \frac{1}{2} g_{ab} (\nabla_c \phi \nabla^c \phi + \mu^2 \phi^2), \quad (8)$$

with the independent non-vanishing components

$$T_{tt} = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}e^{2\Lambda}\phi'^2 + \frac{1}{2}\mu^2\phi^2, \quad (9)$$

$$T_{rt} = \dot{\phi}\phi', \quad (10)$$

$$T_{rr} = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}e^{2\Lambda}\phi'^2 - \frac{1}{2}\mu^2\phi^2, \quad (11)$$

$$T_{\theta\theta} = \sin^{-2} \theta T_{\varphi\varphi} = \frac{1}{2}R^2(\dot{\phi}^2 - e^{2\Lambda}\phi'^2 - \mu^2\phi^2). \quad (12)$$

By defining the auxiliary functions

$$k(t, r) \equiv 1 - e^{2\Lambda}, \quad (13)$$

$$m(t, r) \equiv \frac{1}{2}R(\dot{R}^2 + k), \quad (14)$$

Einstein's equations can be recast in terms of the first derivatives of these two functions:

$$k' = -8\pi R R' (T_{tt} + T_{rr}) - 2R'(\ddot{R} + \dot{\Lambda}\dot{R}), \quad (15)$$

$$\dot{k} = 8\pi R R' T_t^r, \quad (16)$$

$$m' = 4\pi R^2 R' T_{tt} - 4\pi R^2 \dot{R} T_{rt}, \quad (17)$$

$$\dot{m} = 4\pi R^2 R' T_t^r - 4\pi R^2 \dot{R} T_r^r. \quad (18)$$

Since there are only three independent functions to be determined and four equations, only three of these are independent, with the remaining one being a constraint. Since the scalar wave Eq. (6) is implied by the Einstein equations (it follows from the contracted Bianchi identities), we shall take it together with Eqs. (16) and (18) as our complete set. We take Eq. (17) as the constraint equation, since it provides a simple relation between the initial data and the initial mass profile.

To facilitate the resolution of the field equations, we introduce a WKB approximation for the field in the large mass (μ) limit—when the Compton wavelength of the scalar field, μ^{-1} , is much smaller than the radius λ of the spherical region where the field is non-vanishing:

$$\lambda\mu \gg 1. \quad (19)$$

In such a limit, we expect wave-like solutions with slowly-varying (with respect to t) amplitude:

$$\phi(t, r) = \mu^{-1}\Phi(t, r) \cos \mu t. \quad (20)$$

The stress-energy tensor components are then

$$T_{tt} = \frac{1}{2}\Phi^2 - \frac{1}{2}\mu^{-1}\Phi\dot{\Phi} \sin 2\mu t + \frac{1}{4}(\dot{\Phi}^2 + e^{4\Lambda}\Phi'^2) \times (1 + \cos 2\mu t), \quad (21)$$

$$T_{rt} = -\frac{1}{2}\mu^{-1}\Phi\Phi' \sin 2\mu t + \frac{1}{2}\mu^{-2}\dot{\Phi}\Phi'(1 + \cos 2\mu t), \quad (22)$$

$$T_{rr} = -\frac{1}{2}\Phi^2 \cos 2\mu t - \frac{1}{2}\mu^{-1}\Phi\dot{\Phi} \sin 2\mu t + \frac{1}{4}(\dot{\Phi}^2 + e^{4\Lambda}\Phi'^2)(1 + \cos 2\mu t), \quad (23)$$

$$T_{\theta\theta} = \frac{1}{2}R^2[-\Phi^2 \cos 2\mu t - \mu^{-1}\Phi\dot{\Phi} \sin 2\mu t + \frac{1}{2}\mu^{-2}(\dot{\Phi}^2 - e^{4\Lambda}\Phi'^2)(1 + \cos 2\mu t)]. \quad (24)$$

The form of the above equations suggests a trigonometric expansion for $\Phi(t, r)$ of the form

$$\Phi = \Phi_0 + \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mu^{-m} (\Phi_{mn}^c \cos n\mu t + \Phi_{mn}^s \sin n\mu t). \quad (25)$$

The other metric functions are expanded analogously. This expansion in inverse powers of the mass leads to an asymptotic series [14,15] in μ^{-1} , and hence it gives an exact solution in the infinite mass limit and approximate one otherwise. This method is usually referred to as the Lagrange method or the method of averaging [16].

Substitution of the expanded metric functions and their derivatives in Eqs. (6), (14), (16) and (18) fixes the coefficients of all the trigonometric terms at each order in μ^{-1} . Up to $O(\mu^{-2})$ we find that the only non-vanishing terms are:

$$\Phi(t, r) = \Phi_0(t, r) + O(\mu^{-3}), \quad (26)$$

$$m(t, r) = m_0(t, r) + \mu^{-1} m_{12}(t, r) \sin 2\mu t + \mu^{-2} m_{22}(t, r) \cos 2\mu t + O(\mu^{-3}), \quad (27)$$

$$k(t, r) = k_0(t, r) + \mu^{-2} k_{22}(t, r) \cos 2\mu t + O(\mu^{-3}), \quad (28)$$

$$R(t, r) = R_0(t, r) + \mu^{-2} R_{22}(t, r) \cos 2\mu t + O(\mu^{-3}). \quad (29)$$

Taking the first derivatives of the above equations and comparing the leading order terms to the right-hand-side of Eqs. (6), (14), (16) and (18), yields

$$\dot{m}_0 = 0, \quad (30)$$

$$\dot{k}_0 = 0, \quad (31)$$

$$\dot{R}_0^2 = 2m_0 R_0^{-1} - k_0. \quad (32)$$

The leading order terms from Eq. (17) give

$$m'_0 = 2\pi R_0^2 R'_0 \Phi_0^2. \quad (33)$$

The first three equations restrict this class of metrics to the Tolman-Bondi family, which describes the collapse of general inhomogeneous dust. We have thence concluded that sufficiently—in a sense to be defined precisely below—massive scalar fields behave like a dust.

Spherically dust collapse is pressureless and thus always proceeds to a black hole. Therefore, by guaranteeing that the field behaves like dust until the complete formation of the event horizon, we can obtain a sufficient condition for black hole formation from massive scalar field collapse. This criterion is explicitly obtained by examining the next order terms in the WKB approximation. From Eqs. (14), (18) and (16), we have

$$k_{22}(t, r) = 2\pi R'_0(R_0)^{-1}(1 - k_0)\Phi_0\Phi'_0, \quad (34)$$

$$m_{12}(t, r) = \pi R_0^2 \dot{R}_0 \Phi_0^2, \quad (35)$$

$$R_{22}(t, r) = -\frac{\pi}{2} R_0 \Phi_0^2. \quad (36)$$

These terms have the convenient feature of being simple algebraic functions of the known zeroth-order terms and their derivatives. The asymptotic expansion guarantees that the WKB approximation will remain valid whilst all the next order terms remain small compared to the leading order ones:

$$\mu^{-2} \frac{k_{22}}{k_0} \leq \frac{1}{2} \text{ and } \mu^2 \frac{|R_{22}|}{R_0} \leq \frac{1}{2} \text{ and } \mu^{-1} \frac{m_{12}}{m_0} \leq \frac{1}{2}. \quad (37)$$

The value 1/2 was chosen arbitrarily (it has to be smaller than unity!) and changing this value amounts to a change in μ . The WKB approximation will therefore breakdown when one of the above inequalities is no longer satisfied.

The asymptotic expansion implies that the largest correction will come from the m_{12} term, and we have verified this numerically. Thus, the validity of the WKB approximation is given by the condition [where Eqs. (14) and (35) were used]:

$$\dot{R}_0 m'_0 \leq \mu m_0 R'_0, \quad (38)$$

which defines a region on the $t - r$ (hence $t - R$) plane, as shown in Fig.1. The WKB approximation will hold at all points in whose causal past the inequality is always satisfied. On the $t - R$ plane, it is the set of events *outside* the curve that saturates inequality (38), whose past null cone is tangent to it.

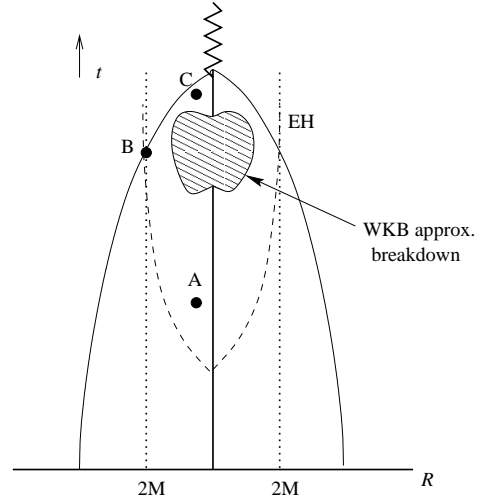


FIG. 1. Spherically symmetric dust collapse in $\{t, R\}$ coordinates. The solid line is the sphere's surface, and the dashed curve that asymptotes $R = 2M$ is the event horizon. The patterned region inside the event horizon is where the WKB condition breaks down. The eikonal approximation holds for events A and B, but not for event C: its past light cone intersects the region where the WKB condition breaks down.

III. THE TOLMAN-BONDI LIMIT

From Eqs. (13), (14), (16), (18), we have (hereafter dropping the '0'-indices for simplicity):

$$ds^2 = -dt^2 + \frac{R'^2(t, r)}{1 - k(r)} dr^2 + R^2(t, r) d\Omega^2, \quad (39)$$

$$\dot{R}^2(t, r) = \frac{2m(r)}{R(t, r)} - k(r), \quad (40)$$

$$\Phi^2(t, r) = \frac{1}{2\pi} \frac{m'}{R^2 R'}. \quad (41)$$

This is the Tolman-Bondi class of solutions. Included in this class are the Schwarzschild metric [$m(r) = \text{const.}$], the Einstein-de Sitter universe [$R(t, r) = a(t)r \propto t^{2/3}r$ and $k(r) = 0$], and the closed Friedmann universe [$k(r) \propto$

r^2]. The metric is written in comoving coordinates and the model describes the pressureless collapse of spherical dust shells. Each shell is labelled by the radial coordinate r and has a surface area $4\pi R^2(t, r)$. In the context of Tolman-Bondi metrics, shell crossings—the overlapping of neighboring shells at finite proper radius—are defined by the locus of events for which $R > 0$ and $R' = 0$. It is clear from Eq. (41) that the energy density diverges at shell crossing singularities. Even though some curvature components may diverge at shell crossings [18], the spacetime is geodesically complete—analytical continuations of the metric can always be found (in a distributional sense) [19]—and thus they are not real physical singularities; rather, they merely signal the intersection of matter flow lines at a given spacelike surface, and thus lack of integrability of the field equations beyond that surface (it is worth pointing out that they also occur in inhomogeneous spherical Newtonian collapse).

Assuming $R' \neq 0$, we can integrate Eq. (40) parametrically to obtain

$$t(\eta, r) = t_0(r) + \frac{m}{k^{3/2}}(\eta + \sin \eta), \quad (42)$$

$$R(\eta, r) = \frac{2m}{k} \cos^2 \frac{\eta}{2}, \quad (43)$$

where $t_0(r)$ is an arbitrary constant of integration (to be fixed by the initial data), and $0 \leq \eta \leq \pi$. The physical initial data consists of a mass distribution $m(r)$ and a velocity profile $\dot{R}(0, r)$. There is gauge freedom for the scaling of r and we shall fix it by requiring that it coincides with the initial area radius:

$$R(0, r) = r. \quad (44)$$

A. Time-symmetric initial data

Considering time-symmetric initial data,

$$\dot{R}(0, r) = 0, \quad (45)$$

implies, from Eq. (40), $k(r) = 2m/r$ and $t_0(r) = 0$. Equations (42) and (43) then simplify to

$$t(\eta, r) = \left(\frac{r^3}{8m} \right)^{\frac{1}{2}} (\eta + \sin \eta), \quad (46)$$

$$R(\eta, r) = r \cos^2 \frac{\eta}{2}. \quad (47)$$

A shell with initial proper area $4\pi r^2$ will thus collapse to vanishing area radius in a (comoving) time

$$t_{\text{coll}}(r) = \pi \sqrt{\frac{r^3}{8m}}. \quad (48)$$

For inhomogeneous mass distributions ($m \neq \text{const.} \times r^3$), different shells become singular at different times; in the

homogeneous case, all the shells collapse to zero area radius at the same time [17].

A particular solution is specified by an initial mass function $m(r)$, given uniquely by the initial density distribution:

$$m(r) = 2\pi \int_0^r \Phi^2(0, \bar{r}) \bar{r}^2 d\bar{r} = 4\pi \int_0^r T_{tt}(0, \bar{r}) \bar{r}^2 d\bar{r}. \quad (49)$$

If the mass function $m(r)$ approaches a constant value M when $r \rightarrow +\infty$, then M is the ADM mass of the spacetime.

IV. THE EINSTEIN-DE SITTER UNIVERSE

The EdeS universe is a particular member of the Friedmann-Robertson-Walker metrics[†],

$$ds^2 = -dt^2 + a^2(t) \left[\frac{dr^2}{1 - Kr^2} + r^2 d\Omega^2 \right], \quad (50)$$

where $K = -1, 0$, or $+1$, for a hyperbolic, flat or spherical spatial geometry, respectively. The Einstein equations are simply

$$\dot{\rho} = -3(\rho + p) \frac{\dot{a}}{a}, \quad (51)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}(\rho + 3p), \quad (52)$$

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi}{3} \rho - \frac{K}{a^2}, \quad (53)$$

where ρ is the proper energy density and p the pressure. For pressureless matter distributions ($p = 0$) and vanishing spatial curvature ($K = 0$)—the EdeS universe—the solution to the Friedmann equation (53) is

$$a(t) = a_0 t^{\frac{2}{3}}, \quad (54)$$

where a_0 is an integration constant. This model describes an open geometry (the $K = 0$ spatial sections are diffeomorphic to \mathbb{R}^{\neq}) in the presence of a constant non-zero energy density distribution. Even though the EdeS spacetime is conformally flat, its causal structure is quite different from asymptotically flat geometries. In particular, and unlike Minkowski or Schwarzschild, past null infinity for EdeS is spacelike, and thus past particle horizons exist [which can be observer (i.e., r) dependent].

The EdeS spacetime is a good approximation to the large scale structure of the universe during a matter dominated phase, when the averaged (over space and

[†]Which are a subclass of the Tolman-Bondi metrics [cf. Eq. (39)], with $R(t, r) = a(t)r$ and $k(r) = Kr^2$.

time) energy density evolves adiabatically and pressures are vanishingly small, as, e.g., immediately after inflation [3–5]. We shall therefore adopt such a metric to model the collapse of overdensity perturbations in the early matter dominated phase that followed inflation.

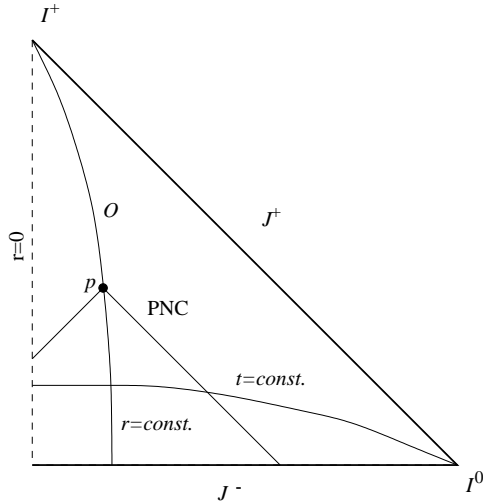


FIG. 2. Conformal diagram for the EdeS spacetime. The dashed vertical line on the left denotes the coordinate (central) singularity. PNC is the past null cone of a timelike observer \mathcal{O} at event p . Worldlines starting from (spacelike) null infinity outside p 's PNC are causally disconnected from \mathcal{O} at p .

A. Dust collapse in the Einstein-de Sitter universe

Our model consists of a general Tolman-Bondi metric and a spherically symmetric matter distribution of finite proper radius $R(0, r) = r = \lambda$, with an overdensity $\zeta(r)$ with respect to the constant background energy density, ρ_c :

$$T_{tt}(0, r) \equiv \rho(r) = \rho_c[1 + \xi(r)], \quad (55)$$

with $\xi(r) = \zeta(r)\Theta(r - \lambda)$, where $\zeta(r)$ is an arbitrary real-valued function and Θ is the unit Heaviside function. For $r \gg \lambda$, the metric asymptotically approaches the EdeS metric and the finite background energy density ρ_c implies the existence of a cosmological horizon with proper radius $H^{-1} = (8\pi\rho_c/3)^{-1/2}$. Clearly, the existence of such a horizon (or, equivalently, expanding exterior geometry) precludes the use of time-symmetric initial data.

Let us then define the generalized Hubble parameter by

$$H(t, r) \equiv \frac{\dot{R}}{R}. \quad (56)$$

Then, from Eqs. (40), (42), (43), we have

$$H_0(r) = \frac{2m}{r^3} \sin^2 \frac{\eta_0}{2}, \quad (57)$$

$$k(r) = \frac{2m}{r} \cos^2 \frac{\eta_0}{2} = \frac{2m}{r} \left(1 - \frac{H_0^2 r^3}{2m}\right), \quad (58)$$

$$t_0(r) = -\frac{m}{k^{3/2}}[\eta_0(r) + \sin \eta_0(r)], \quad (59)$$

where $H_0(r) \equiv H(0, r) = \dot{R}(0, r)/r$. Initial data consists of a mass profile $m(r)$ and an initial Hubble parameter $H_0(r)$. We shall consider here the case of an EdeS universe with initial expansion rate

$$H_0^2 = \frac{8\pi}{3} \rho_c. \quad (60)$$

From Eqs. (49) and (55), the mass function is then

$$m(r) = 4\pi \int_0^r T_{tt}(0, x) x^2 dx = \frac{4\pi}{3} \rho_c r^3 (1 + \bar{\xi}) \\ = M \left(\frac{r}{\lambda}\right)^3, \quad (61)$$

where $\bar{\xi}$ is the volume average of $\xi(r)$ and M is the mass inside a sphere of proper area $4\pi R^2(t, \lambda)$.

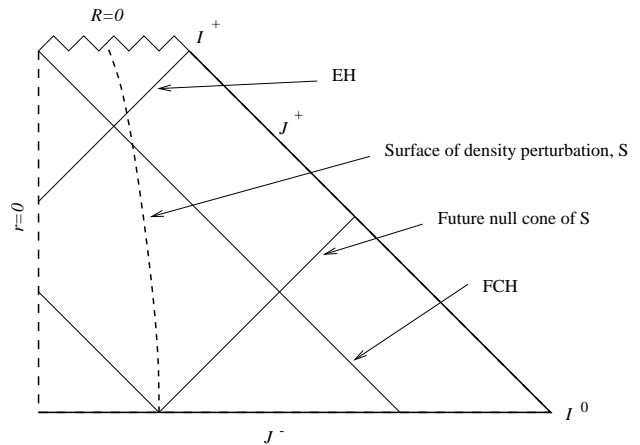


FIG. 3. Conformal diagram for spherical dust collapse in the EdeS universe. As the overdensity collapses (dashed curve), it first crosses the future cosmological horizon (FCH) of an $r = 0$ observer, and, at a later time, the black hole event horizon, eventually ending up at the $R = 0$ central singularity. Events outside the FCH will never be seen by the $r = 0$ observer. Events outside the future null cone of S (the overdensity's spacelike 2-surface) will never be influenced by the overdensity perturbation.

V. CONDITIONS FOR BLACK HOLE FORMATION

We want to impose the condition that the stationary phase approximation holds [cf. Eq. (38)]:

$$\dot{R}m' < \mu R'm, \quad (62)$$

where

$$R' = R \left(\frac{m'}{m} - \frac{k'}{k} \right) + \tan \left[t'_0 \sqrt{k} + \frac{m}{k} \gamma (\eta + \sin \eta) \right], \quad (63)$$

$$\dot{R} = -\sqrt{k} \tan \frac{\eta}{2}, \quad (64)$$

$$t'_0 = -\gamma t_0 - \frac{m}{k^{3/2}} \cot \frac{\eta}{2} \cos^2 \frac{\eta}{2} \left(\gamma + \frac{k'}{2k} - \frac{1}{r} \right), \quad (65)$$

$$\gamma(r) \equiv \frac{m'}{m} - \frac{3}{2} \frac{k'}{k}. \quad (66)$$

For simplicity, we consider the energy density distribution inside the overdense region to be constant and hence, from Eqs. (61), (65) and (66), we have $\gamma(r) = t'_0(r) = 0$. The relevant WKB condition (62) becomes then

$$3p^3 \sqrt{\frac{2M}{\lambda^3}} < \mu \cos^2 \frac{\eta}{2} \cot \frac{\eta}{2}, \quad (67)$$

where

$$p(\bar{\xi}) \equiv (1 + \bar{\xi})^{-\frac{1}{2}} \in [0, 1] \quad (68)$$

is a monotonically increasing function of the magnitude of the volume-averaged overdensity, satisfying $p(0) = 0$ and $\lim_{\bar{\xi} \rightarrow +\infty} p = 1$.

Outside the overdensity region, for $r > \lambda$, the Tolman-Bondi solution is exact: the EdeS metric is a particular case, trivially obtained from the general Tolman-Bondi form by setting $R(t, r) = (t/t_i)^{2/3} r$ and $k(r) = 0$, where t_i is the initial slice. Therefore it suffices to enforce the above inequality (67) for $0 < r \leq \lambda$, to ensure that the WKB approximation will hold for all $r > 0$. If all the shells with $0 < r \leq \lambda$ cross the event horizon *before* the WKB breaks down, then we can conclude that a black hole has formed.

In order to find the loci of events defined by the WKB condition and place them wholly inside the event horizon (EH), we need to (i) find a suitable parameterization for the curve that bounds the region defined by Eq. (67), and (ii) solve for the EH with the same parametrization for the null generators.

The natural such parameterization for the curve that saturates inequality (67) is $\eta_*(r)$ [since η and p (through $\bar{\xi}$) depend solely on r], given by:

$$3p^3 \sqrt{\frac{2M}{\lambda^3}} = \mu \cos^2 \frac{\eta_*}{2} \cot \frac{\eta_*}{2}. \quad (69)$$

From Eqs. (39) and (42), we can parameterize the null generators by η_{EH} , obtained by integration of the outgoing radial null geodesics equation:

$$\frac{d\eta_{\text{EH}}}{dr} = \left\{ R'(k^{-1} - 1)^{-\frac{1}{2}} - \left[t'_0 \sqrt{k} + \frac{m}{k} \gamma (\eta + \sin \eta) \right] \right\} R^{-1}. \quad (70)$$

For $r \leq \lambda$, this simplifies to

$$\frac{d\eta_{\text{EH}}}{dr} = \left(\frac{\lambda^3}{2Mp^2} - r^2 \right)^{-\frac{1}{2}}, \quad (71)$$

which is trivially integrated to give

$$\eta_{\text{EH}}(r \leq \lambda) = \eta_{\text{EH}}(0) + \sin^{-1} p \sqrt{\frac{2M}{\lambda}} \frac{r}{\lambda}. \quad (72)$$

Outside the overdense region, $\gamma \neq 0$ and $t'_0 \neq 0$, which makes it impossible to integrate the radial null geodesics equation analytically. Since the exterior geometry is not asymptotically flat, we cannot use Birkhoff's theorem to place the EH at $R = 2M$. We can, however, solve for the apparent horizon (AH)—the outer boundary of a closed spacelike surface whose normal null geodesics have vanishing divergence—at any given $t = \text{const.}$ hypersurface. In spherical symmetry, the equation for the AH is simply the requirement that a surface $R(t, r) = \text{const.}$ becomes null:

$$R^a R_{,a} = 0. \quad (73)$$

From Eqs. (39) and (40), we have then

$$R(t_{\text{AH}}, r) = 2m(r), \quad (74)$$

which, via Eq. (42), defines a curve $\eta_{\text{AH}}(r)$, given by

$$\eta_{\text{AH}}(r) = 2 \cos^{-1} \frac{r}{\lambda} \sqrt{\frac{2M}{\lambda}} p. \quad (75)$$

Since the AH must be contained or coincide with the (spacelike section of the) EH, we must have $\eta_{\text{AH}}(\lambda) \geq \eta_{\text{EH}}(\lambda)$, which gives an upper bound for $\eta_{\text{EH}}(0) \equiv \eta_0$:

$$\eta_0 \leq 2 \cos^{-1} \sqrt{\frac{2M}{\lambda}} p - \sin^{-1} \sqrt{\frac{2M}{\lambda}} p. \quad (76)$$

This upper bound on η_0 guarantees that the approximate EH thus estimated coincides or lies inside the actual EH; hence, if the WKB approximation holds inside the former, it will obviously hold inside the latter.

We now need to guarantee that the breakdown domain of the WKB approximation is placed entirely inside the EH. From Eq. (69), η_* is implicitly defined by an expression of the form $f[\eta_*] = \frac{1}{3} p^{-3} \sqrt{\frac{\lambda^3}{2M}}$, where f is a monotonically increasing functional of η_* ; therefore, $\eta_* > \eta_{\text{EH}}(\lambda) \Rightarrow f[\eta_*] > f[\eta_{\text{EH}}(\lambda)]$. This enables a sufficient condition for the WKB approximation to hold (inside the true EH), by imposing a lower bound on η_* :

$$\eta_* > \eta_{\text{EH}}(\lambda) \geq 2 \cos^{-1} \sqrt{\frac{2M}{\lambda}} p. \quad (77)$$

From the above equation, together with Eq. (67), we finally obtain

$$\mu M > \frac{3}{2} \sqrt{1 - \frac{2M}{\lambda}} p^2. \quad (78)$$

Matter configurations with parameters satisfying the above inequality will collapse to form a black hole. This is a sufficient, but not necessary, criterion, due to both the adiabatic ansatz ($\mu\lambda \gg 1$) and the approximation for the EH. There are two limits that should be pointed out:

In the limit of very large overdensities, the inequality approaches the one for the asymptotically flat case [8], as expected. The simple physical reason is that the ratio $\rho_c/\rho_{\text{pert.}} \ll 1$, and therefore having a vanishing or finite background energy density becomes irrelevant; the overdense region does not “feel” the exterior geometry. This can be heuristically understood as follows: The typical cosmological expansion rate is $(\dot{R}/R)_{\text{exp}} \sim \rho_c^{-1/2}$, whereas the collapse rate is $(\dot{R}/R)_{\text{coll}} \sim -[(1+\bar{\xi})\rho_c]^{-1/2}$. Then, $\frac{\tau_{\text{exp}}}{\tau_{\text{coll}}} = \left| \left(\frac{\dot{R}}{R} \right)_{\text{coll}} \right| / \left(\frac{\dot{R}}{R} \right)_{\text{exp}} \sim (1+\bar{\xi})^{1/2} \gg 1$: the expansion that the overdensity would have experienced during a (proper) time t_{coll} , had it not collapsed during that time, is negligible compared to its initial physical radius.

When the overdensity vanishes, although the right-hand-side of the inequality approaches $3/2$, this result is meaningless because it was based on an estimate for the EH location that becomes totally unreliable: $\eta_{\text{EH}}(0) = \pi$, when $p = 0$. This is trivial, since $p = 0$ corresponds to an unperturbed EdeS universe. For finite $p \ll 1$, provided the spatial extent of the overdensity is sufficiently large, $\lambda > 2Mp^2$, the inequality is still valid.

We can also look at this condition in terms of the cosmological horizon, H_0^{-1} . Noting that $2M = H_0^2 \lambda^3 (1 - p^2)^{-1}$ and recalling that $p \equiv (1 + \bar{\xi})^{-1/2}$, we can rewrite the condition for black hole formation as

$$\mu M > \frac{3}{2} \sqrt{1 - (H_0 \lambda)^2 \bar{\xi}}. \quad (79)$$

The two relevant limits are now:

The length scale λ is much smaller than the horizon radius H_0^{-1} , and $\bar{\xi}$ can be sufficiently large to recover the $p \rightarrow 1$ limit and thus the asymptotically flat results. This limit also provides a useful self-consistency check for the validity of the WKB approximation. Let us take $H_0 \lambda \ll 1$ and $\bar{\xi} \sim O(1)$; then, the inequality becomes $H_0^2 \lambda^3 \mu \gtrsim \frac{3}{2}$. Since, by hypothesis, $H_0 \lambda \ll 1$, the necessary and sufficient condition to enforce the inequality is $\mu \lambda \gg 1$, which is precisely the adiabatic condition on which the whole argument is based. Thus, we confirm that if the WKB approximation is valid initially, it will remain valid up to the formation of the EH.

If $\bar{\xi} > (H_0 \lambda)^{-2}$, then the right-hand-side of the inequality becomes imaginary and the approximation breaks down. Physically, this corresponds to an overdensity that becomes so large that a spacelike hypersurface would curve upon itself and form a disjoint spatial universe, which does not correspond to a black hole [20]. Provided $\bar{\xi} < (H_0 \lambda)^{-2} \lesssim 1$, this limit is equivalent to the $p \ll 1$ limit previously described, and thus still valid. From an astrophysical viewpoint, this is the most interesting limit, as it constrains the density perturbations at horizon crossing.

VI. CONSTRAINING THE PRIMORDIAL BLACK HOLE SPECTRUM

Even though inflation is still a scenario and not yet a paradigm for cosmology, its basic predictions appear fairly robust in that not only they solve the (“naturalness” and “fine-tuning”) shortcomings of the “old” standard Big Bang model, but they also present a viable physical mechanism for the primordial density fluctuations that seeded the subsequent structure formation.

As previously mentioned, several inflationary models predict the existence of a period dominated by the energy density of a real massive scalar field, just after the end of inflation, hereafter denoted by t_{end} . During this stage, the field undergoes coherent oscillations, behaving effectively like a dust, from t_{end} until t_{RH} (“reheating”), when it rapidly decays into relativistic particles [21]. Any density perturbations crossing the horizon during this dust-like phase will inevitably (there is no pressure) collapse to form black holes, provided they are sufficiently spherical, as aspherical growth precludes the gravitational collapse of the inhomogeneities [22]. Clearly, it is during this intermediate matter-dominated phase that black hole formation via gravitational collapse of (spherical) overdensities is the most abundant. We should therefore expect important constraints on inflationary scenarios (e.g., on the decay width of the scalar field, Γ) from the overproduction of primordial black holes (PBH) during this period.

The present process will only be relevant for black hole formation if the time interval, Γ^{-1} , after which the field decays into radiation is larger than the collapse time to the central singularity. We note that it is sufficient to require that the scalar field dominates just until the overdense region has collapsed through its Schwarzschild radius, since a subsequent decay into radiation (before vanishing proper area is reached) would not destroy the already formed black hole; furthermore, the proper time to reach the formation of black hole horizons (i.e., the proper time elapsed between horizon crossing and the

emission of the null generators) is much larger than the time it would take for a black-hole-sized configuration to collapse to a singularity.

Since spherical symmetry is assumed throughout, the necessary and sufficient conditions for black hole formation from a massive scalar field configuration with parameters $\{\mu, M, \lambda\}$ are that black holes:

$$\text{have time to form: } t_{\text{HC}} + t_{\text{coll}} < t_{\text{RH}}, \quad (80)$$

$$\text{will form: } \mu M > \frac{3}{2} \sqrt{1 - (H_0 \lambda)^2 \bar{\xi}}, \quad (81)$$

where in the first inequality, the subscripts refer to “horizon crossing”, “collapse” and “reheating”, respectively. Because the synchronous gauge is being used, the three timescales can be trivially compared, thus justifying inequality (80). Since we are considering homogeneous density distributions, the proper time for the collapse of the overdense region will depend solely on its density. Setting $\eta = \pi$, from Eqs. (42), (43) and (57)-(59), we obtain

$$t_{\text{coll}} = \frac{1}{2} H_0^{-1} \bar{\xi}^{-\frac{3}{2}} (1 + \bar{\xi}) (\pi - \sin^{-1} \chi - \chi), \quad (82)$$

$$\chi \equiv 2\sqrt{\bar{\xi}}(1 + \bar{\xi})^{-1}. \quad (83)$$

For large overdensities, $\bar{\xi} \gg 1$, $t_{\text{coll}} \sim \frac{1}{2} \pi H_0^{-1} \bar{\xi}^{-\frac{3}{2}}$, whereas in the limit of small perturbations, $\bar{\xi} \ll 1$, the collapse time is

$$t_{\text{coll}} \approx \frac{\pi}{2} H_0^{-1} \bar{\xi}^{-\frac{3}{2}}. \quad (84)$$

In the EdeS universe, $H_0 = \frac{2}{3} t^{-1} = \sqrt{8\pi\rho_c/3}$. At horizon crossing, $H\lambda = 1$; thus

$$t_{\text{HC}} = \frac{2}{3} \lambda = \frac{2}{3} H_0^{-1}. \quad (85)$$

With Eq. (84), we can rewrite Eq. (80) as (dropping the bar for simplicity)

$$\xi > \left(\frac{\pi}{2}\right)^{\frac{2}{3}} \left(H_0 t_{\text{RH}} - \frac{2}{3}\right)^{-\frac{2}{3}} \equiv \xi_{\text{min}}^{(1)}. \quad (86)$$

That is, at a given slice $t = \frac{2}{3} H_0^{-1}$, there will be a minimum fractional density perturbation at horizon crossing, below which the scalar field will decay into radiation before the overdensity has time to undergo complete gravitational collapse.

Let us now examine the WKB condition at horizon crossing. The mass of a perturbation with fractional overdensity ξ , at horizon crossing, is

$$\begin{aligned} M &= \frac{4}{3} \pi \rho_c (1 + \xi) \lambda^3 = \frac{1}{2} H_0^2 (1 + \xi) \lambda^3 \\ &= \frac{1}{2} H_0^{-1} (1 + \xi) \simeq \frac{1}{2} H_0^{-1}. \end{aligned} \quad (87)$$

Equation (81) becomes then

$$\mu M > \frac{3}{2} \sqrt{1 - \xi}, \quad (88)$$

which places another lower limit on ξ :

$$\xi > 1 - \frac{4}{9} \mu^2 M^2 \equiv \xi_{\text{min}}^{(2)}. \quad (89)$$

We have then two independent lower bounds on ξ , that tell us, respectively, that black holes *can* and *will* form. If $\mu M > \frac{3}{2}$, then the dust approximation remains reliable and Eq. (89) is always satisfied. In this case, Eq. (86) becomes the relevant constraint, and the crucial quantity is $H_0 t_{\text{RH}}$, which, although model dependent (through t_{RH} , or, equivalently, Γ^{-1} , for a given horizon crossing time), is expected to be large, $H_0 t_{\text{RH}} > 10^6$ [21].

If $\mu M < 3/2$, then the dust approximation may become unreliable and the fractional density perturbations are of order unity, therefore departing from the linear regime and being heavily suppressed in the initial perturbation spectrum, in which case the dust-like phase is irrelevant for the production of PBH from the gravitational collapse of massive scalar fields.

We shall then consider only configurations satisfying $\mu M > 3/2$. Let us first rewrite the density contrast in more familiar notation:

$$\xi = \frac{\rho}{\rho_c} - 1 = \frac{\Delta\rho}{\rho_c} \equiv \delta, \quad (90)$$

where δ is to be evaluated at horizon crossing.

For the spectrum of density fluctuations crossing the horizon, we consider a Gaussian probability distribution (as predicted by inflation [21]):

$$P(\delta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\delta^2}{2\sigma^2}\right), \quad (91)$$

where σ is the root-mean-square amplitude of the density fluctuations. Provided we restrict ourselves to linear perturbations, the Fourier modes will evolve independently, thereby preserving Gaussianity. (Non-Gaussian effects arise if $\delta \gg \sigma$, when linearity is spoiled by mode-mode coupling, but their influence on the initial PBH mass function for a dust-dominated era has been shown to be negligible [23]).

The quantity of interest in the constraining of PBH production is the mass fraction of the universe that collapsed to a black hole during the intermediate dust-like epoch:

$$\beta \equiv \frac{\rho_{\text{PBH}}}{\rho_c} = \int_{\xi_{\text{min}}}^1 P(\delta) d\delta, \quad (92)$$

where ρ_{PBH} is the initial (i.e., formed during that epoch) energy density of primordial black holes, and ρ_c is the background energy density. ξ_{min} is given by Eq. (86) and the upper integration limit is imposed by linearity.

Expressed in this way, β is model dependent, since ξ_{\min} depends on H_0 and t_{RH} (equivalently, Γ^{-1}).

β can be observationally constrained via (i) the requirement that the present density fraction of PBH's, $\Omega_{\text{PBH}} < 1$; and (ii) limits from the Hawking radiation, from upper bounds on the PBH population at subsequent epochs, e.g., baryogenesis and nucleosynthesis. These constraints on β will lead to constraints on the cosmological parameters β is implicitly dependent on, namely the decay width of the scalar field and the time for horizon crossing. Work in this direction is currently underway and will be presented in detail elsewhere [7].

VII. CONCLUSIONS

We have shown that there is an adiabatic regime ($\lambda\mu \gg 1$), in which massive scalar fields in an asymptotically EdeS background behave like a dust and therefore inevitably collapse to form black holes. Enforcing the adiabaticity condition to second order in μ^{-1} lead to a simple analytical criterion to be obeyed by the initial data to collapse to black holes. This criterion is valid for spherical density perturbations of *any* finite physical radius λ ; in particular, they can equal the horizon size, $H\lambda = 1$. Perturbations with spatial radius larger than the horizon radius are physically uninteresting because (i) the fractional density perturbation required to satisfy the WKB condition is very small, $\xi \ll (H\lambda)^{-2}$, and (ii) they are causally disconnected from the observable universe.

Immediately after inflation, there can exist a short period dominated by the energy density of a real massive scalar field [3–5]. By considering spherical density perturbations that re-enter the horizon during this epoch, we obtained a lower limit on the fractional density perturbation (at horizon crossing) that would allow black holes to form before the field decays into relativistic particles. This lower limit can then be used to compute the initial mass function of PBH at the early matter-dominated phase. Constraints on the PBH mass spectrum from its subsequent evolution will ultimately lead to constraints on some free parameters of particle astrophysics models [7].

ACKNOWLEDGMENTS

I am grateful to Bernard Carr, Ian Moss, Kip Thorne and Michele Vallisneri for helpful discussions. This work was supported by F.C.T. (Portugal) Grant PRAXIS XXI-BPD-16301-98, and by NSF Grant AST-9731698.

- [1] J. F. Donoghue, E. Golowich and B. R. Holsten, *Dynamics of the Standard Model* (Cambridge University Press: Cambridge, 1992).
- [2] A. H. Guth, “Inflationary Models and Connections to Particle Physics”, astro-ph/0002188.
- [3] S. W. Hawking and I. G. Moss, Phys. Lett. **B110**, 35 (1982).
- [4] A. D. Linde, Phys. Lett. **B129**, 177 (1982).
- [5] A. Albrecht and P. J. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982).
- [6] Yu. M. Khlopov, B. A. Malomed and Ya. B. Zel’dovich, Mon. Not. R. Astron. Soc. **215**, 517 (1985).
- [7] S. M. C. V. Gonçalves, in preparation.
- [8] S. M. C. V. Gonçalves and I. G. Moss, Class. Quantum Grav. **14**, 2607 (1997).
- [9] G. Lemaitre, Ann. Soc. Sci. Bruxelles I **A53**, 51 (1933).
- [10] R. C. Tolman, Proc. Nat. Acad. Sci. USA **20**, 410 (1934).
- [11] B. Datt, Z. Phys. **108**, 314 (1938).
- [12] H. Bondi, Mon. Not. Astron. Soc. **107**, 343 (1947).
- [13] I. H. Dwivedi and P. S. Joshi, Class. Quantum Grav. **9**, L69 (1992).
- [14] J. L. Lagrange, *Mecanique Analytique* (originally published in 1788; Edition Albert Blanchard: Paris, 1965).
- [15] N. N. Bogoliubov and Yu. A. Mitropolski, *Asymptotic Methods in the Theory of Nonlinear Oscillations* (Gordon and Breach: New York, 1981).
- [16] F. Vershult, *Nonlinear Differential Equations and Dynamical Systems* (Springer-Verlag: Berlin, 1991).
- [17] This also happens in the Newtonian collapse of homogeneous spherical shells, and it is due to the homogeneity of the matter distribution and its spherical symmetry: homogeneity implies that $m(r) \propto r^3$ and spherical symmetry guarantees that the gravitational force per unit mass at radius r is $F \propto mr^{-2}$. Since the equation of motion for a unit test mass is $\ddot{r} = F \propto r$, it follows, from dimensional analysis alone, that the collapse time will be *independent* of the radial coordinate r . Outer shells are more accelerated (because the mass inside them is larger) than inner ones, and thus collapse faster to travel a larger distance in *exactly* the same time needed for inner, less accelerated shells.
- [18] P. Yodzis, H. J. Seifert, and H. Muller zum Hagen, Comm. Math. Phys. **34**, 135 (1973).
- [19] A. Papapetrou and A. Hamoui, Ann. Inst. Henri Poincaré **VI**, 343 (1967).
- [20] B. J. Carr and S. W. Hawking, Mon. Not. R. Astron. Soc. **168**, 399 (1974).
- [21] E. Kolb and M. S. Turner, *The Early Universe* (Addison-Wesley Publishing: Reading, Massachusetts, 1994).
- [22] M. Yu. Khlopov and A. G. Polnarev, Phys. Lett. **B97**, 383 (1980).
- [23] J. S. Bullock and J. R. Primack, Phys. Rev. D. **55**, 7423 (1997).